

# Modular design increases energy efficiency

**Robert Roe speaks to Bill Thigpen, advanced computing branch chief for the Nasa Advanced Supercomputing Division**



Nasa has been experimenting with a modular build for its latest supercomputer, Electra. First installed in 2016 and set-up for full operation in 2017, the system is based on SGI E-cell design. The aim of the supercomputer is to deliver higher energy efficiency for Nasa's computing operations, while still providing the necessary computational power for Nasa's science and engineering workloads.

At the time of the first announcement of the modular design, Bill Thigpen, advanced computing branch chief for the Nasa Advanced Supercomputing (NAS) Division, said: 'If you look at any of the top systems, they are all drawing multiple megawatts of power. They are taking somewhere in the region of a 33 to 50 per cent of that power, just for cooling.

'These systems draw a large amount of power, but if we can save something on the order of a third of the power by doing things in a smarter way, I think it is our duty to do that. We should be good stewards of the earth.'

NAS operates Nasa's High-End Computing Capability (HECC) Project, funded through the agency's High-End Computing Program and its Strategic Capabilities Assets Program. HECC supports more than 1,200 users from around the US, with more than 500 projects at any one time.

The NAS Division is part of the Exploration Technology Directorate at Ames Research Center. The directorate's mission is to create innovative and reliable technologies for Nasa missions.

The first iteration of Electra was a 16-rack, 1,152-node cluster from SGI, with a 1.2Pflop peak performance. The system was listed as delivering 1.09Pflop/s LINPACK rating on the November 2016 TOP500 list – putting it at 96 on that list.

Housed in two containers that comprise the Modular Supercomputing Facility (MSF) at Nasa's Ames Research Center, the initial modules for Electra were expanded over the past year to provide researchers with more computing power to support Nasa's scientific and engineering projects. Systems engineers at the NAS facility added 1,152 Intel Xeon Gold 6148 Skylake nodes last year, doubling the number of Electra's Skylake nodes to 2,304 and fully populating the system's second module. Combined with the first module's 1,152 Xeon E5-2680v4

**"These systems draw a large amount of power, but if we can save something on the order of a third of the power by doing things in a smarter way, I think it is our duty to do that"**

(Broadwell) nodes, the expansion brings Electra's theoretical peak performance to 8.32Pflops.

After the upgrade, Electra is the 12th most powerful computer in the US and ranks 33rd worldwide on the November 2018 TOP500 list.

The system reached ninth in the US and 24th in the world on the latest High Performance Computing Gradient (HPCG) benchmark list.

By expanding the compute power available to Nasa scientists and engineers, Electra has already made a large impact on the Nasa's research objectives. Researchers at the NAS facility are now using the supercomputer for a wide array of research projects. This includes Nasa

aerospace engineer Patricia Ventura Diaz, who is working on urban air mobility with high-fidelity simulations of autonomous rotary-wing concept vehicles; Jordan Angel, who researches the feasibility of electric aircraft designs and quiet supersonic technology; and Michael Aftosmis, who helps assess the risk posed by asteroids entering Earth's atmosphere by running thousands of potential impact scenarios on Electra.

This year NAS are finalising the upgrade to the modular design of Electra. The system is now being expanded from the initial two modules with an additional eight E-cells using Skylake processors from Intel.

With a year of data collected from the operational system, the NAS team can now understand just how this project has changed the requirements for cooling and water usage.

'Electra is now an 8.3Pflop computer. There are eight E-cells in the second module which is about 7Pflops, which is all Skylake CPUs, and then we have the original Broadwells that are in the original module.

'One of the things that I was thrilled about, is that we now have a year of data. And so in that year we actually did really well. The PUE ended up being 1.031 for the combined system. That is October 2017 to September 2018.'

One approach to solving this problem is being developed by NAS as part of the HECC project. Nasa hopes to solve some of its future power issues by developing a modular HPC cluster using energy-efficient technology and making use of evaporative cooling and the local climate to remove the need for much of the water and power used to cool the system.

Thigpen noted that the amount of water and electricity used in the system had been massively reduced, compared to previous generations of supercomputers. 'We calculated how much water we would have used in the building based on the electrical load, and what it would take to get that heat out. The calculated number was 9.2 million gallons. What we actually used in the facility was 128,000 gallons, so that was

The Crab Nebula

a saving of around 96 per cent,' noted Thigpen.

'When we look at the electricity that we used to cool the system in the building, it would have been 2.4 million kilowatt hours and we used 221 thousand kilowatt hours, so we saved almost 91 per cent on energy costs.'

'That has gone really well, so the new site is under construction right now. We have got the power running to the site. We had to expand the substation and we are bringing more power over to the site than we think we will need,' added Thigpen.

The success of the modular system architecture has led NAS to expand the system systematically over the next few years. While Thigpen and his colleagues expect the PUE numbers to be largely the same, the number of E-cells per module will increase, delivering more computational power from each module.

'We have plenty of power to meet more than I could ever think we will ever install but it's much simpler to bring the power over to the area, than to do it after the fact. We have a primary and secondary

**"We want to ensure that what we deploy matches the demand, so as demand in that area grows we will increase the resources accordingly"**

infrastructure stage and the primary stage is basically done now, and the secondary stage should start as soon as the permits are cleared for the centre.

'What that means is that all the power, water, roads, fencing, all of that type of work is done, and so all that is left is pouring the concrete pad and installing the first module on the site. Each module will be around 50 per cent bigger than our second module, module 2, which holds eight E-cells; the module on the new site will hold 12 E-cells.'

The new modules are expected to go into operation sometime after March this year. The modules already have the infrastructure in place and will share the same file system used in the current

Electra system. The site will eventually hold 12 modules and four data modules. Thigpen noted that a module full of Skylake processors can deliver as much as 10.5Pflops per module. This can be increased further through the use of accelerators, but NAS is still trying to determine how beneficial accelerator technologies will be for its existing user base.

The current plan is to install a small number of GPUs in the Pleiades system, so that they can be tested with current workloads run at the NAS facility.

'In the original deployment we are not planning on having accelerators. We are buying some accelerators for our facility and are in the process of placing an order where we will place 16 nodes, and 14 of those nodes will have four V100's and two of those nodes will have eight V100's,' said Thigpen.

'We would look at putting the right mix of accelerators. We want to ensure that what we deploy matches the demand, so as demand in that area grows we will increase the resources accordingly,' he concluded. ■